

Q68HV1\_MOUSE  
 ID Q68HV1\_MOUSE PRELIMINARY; PRT; 1925 AA.  
 AC Q68HV1;  
 DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.  
 DT 11-OCT-2004, sequence version 1.  
 DT 13-JUN-2006, entry version 11.  
 DE Plexin D1.  
 GN Name=Plxnd1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muroidae; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=BALB/c;  
 RA Duke-Cohan J.S., Ahmed W., Reinherz E.L.;  
 RT "Identification of mouse plexin D1."  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AY688678; AAT99561.1; -, mRNA.  
 DR UniGene; Mm.3085; -.  
 DR Ensembl; ENSMUSG00000030123; Mus musculus.  
 DR MGI; MGI:2154244; Plxnd1.  
 DR RZPD-ProtExp; IOM17766; -.  
 DR GO; GO:0005886; C:plasma membrane; IC.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR GO; GO:0017154; F:semaphorin receptor activity; IPI.  
 DR GO; GO:0001569; P:patterning of blood vessels; IMP.  
 DR InterPro; IPR002909; IPT\_TIG\_rcpt.  
 DR InterPro; IPR003659; Plexin-like.  
 DR InterPro; IPR013548; Plexin\_cytopl.  
 DR InterPro; IPR002165; Plexin\_repeat.  
 DR InterPro; IPR008936; Rho\_GAP.  
 DR InterPro; IPR001627; Sema.  
 DR Pfam; PF08337; Plexin\_cytopl; 1.  
 DR Pfam; PF01437; PSI; 2.  
 DR Pfam; PF01403; Sema; 1.  
 DR Pfam; PF01833; TIG; 3.  
 DR SMART; SM00429; IPT; 3.  
 DR SMART; SM00423; PSI; 3.  
 DR SMART; SM00630; Sema; 1.  
 DR PROSITE; PS51004; SEMA; 1.  
 SQ SEQUENCE 1925 AA; 211638 MW; 53E71ADB05A78FC CRC64;  
 Query Match 100.0%; Score 9181; DB 2; Length 1925;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	180		239
Qy	61	IAIRSLDARGDLAKLFTFDLNPSSDDNLIKQGAKEQHKLGFRVAFVRAFLHPAVPPHSAQPYA	120
Db	240	IAIRSLDARGDLAKLFTFDLNPSSDDNLIKQGAKEQHKLGFRVAFVRAFLHPAVPPHSAQPYA	299
Qy	121	YLALNSEARAGDKDSQARSLLARICLPRGAGGDAKKLTESYIQLGLQCAGGAGRGDLYSR	180
Db	300	YLALNSEARAGDKDSQARSLLARICLPRGAGGDAKKLTESYIQLGLQCAGGAGRGDLYSR	359
Qy	181	LVSVPFPAEQFFAVFERPQAGPARNAPALCAFRFDDVQAAIRAARTACFVEPAPDVVA	240
Db	360	LVSVPFPAEQFFAVFERPQAGPARNAPALCAFRFDDVQAAIRAARTACFVEPAPDVVA	419
Qy	241	VLDSPVQGTGPACESKRNIQLQPEQLDCGAHLQHPLTILQPLRASPVPFRAPGLTAVAVA	300
Db	420	VLDSPVQGTGPACESKRNIQLQPEQLDCGAHLQHPLTILQPLRASPVPFRAPGLTAVAVA	479
Qy	301	SANNYTAVFLGTATGRLLKISLNEISMVQVSRRLVTAYGEPVHHVMQFDPMPGYLYLMT	360
Db	480	SANNYTAVFLGTATGRLLKISLNEISMVQVSRRLVTAYGEPVHHVMQFDPMPGYLYLMT	539
Qy	361	SHQMARVKVAACEVHSTCGDCVGAADAYCGWCTLETRCTLQDDCTNSSQPHFWTSASEGP	420
Db	540	SHQMARVKVAACEVHSTCGDCVGAADAYCGWCTLETRCTLQDDCTNSSQPHFWTSASEGP	599
Qy	421	SRCPAMTVLPSEIDVHRDYGTMILQISGSLPSSGEMACDYGNGVRTVARVPGPAYDHQ	480
Db	600	SRCPAMTVLPSEIDVHRDYGTMILQISGSLPSSGEMACDYGNGVRTVARVPGPAYDHQ	659
Qy	481	IAYCNLLPRAQFSPFPAQGDHVTIVMSVRVKGNHIVSANFTIYDCSRIGQVYPHTACTSC	540
Db	660	IAYCNLLPRAQFSPFPAQGDHVTIVMSVRVKGNHIVSANFTIYDCSRIGQVYPHTACTSC	719
Qy	541	LSTQWPCSWCIQLHSCVSNQSQQDSNPNTSPQDCPQILPSPLAPVPTGGSQDILVPLTK	600
Db	720	LSTQWPCSWCIQLHSCVSNQSQQDSNPNTSPQDCPQILPSPLAPVPTGGSQDILVPLTK	779
Qy	601	ATFFHGSSLECSFGLSESFVAVWANNLSVRCNQVVLHTTQKSQVFPPLSKLKGPPDRFLD	660
Db	780	ATFFHGSSLECSFGLSESFVAVWANNLSVRCNQVVLHTTQKSQVFPPLSKLKGPPDRFLD	839
Qy	661	SPNPMTVVYVNCAMGSPDCSQCLGREDLGHLCVWNGDGRCLRGPLQPLPGTCFAPEIRAIE	720
Db	840	SPNPMTVVYVNCAMGSPDCSQCLGREDLGHLCVWNGDGRCLRGPLQPLPGTCFAPEIRAIE	899
Qy	721	PLSGPLDGGTLLTIRGRNLGRRLSDVAGHVWIGSVACEPLADRYTVSEEIVCATGPAAGA	780
Db	900	PLSGPLDGGTLLTIRGRNLGRRLSDVAGHVWIGSVACEPLADRYTVSEEIVCATGPAAGA	959
Qy	781	FSDVTVTVNVSKEGRSREQFSYVLPVHSLPESMGPKAGGTRITIHGSDNLVNGSMLQVLVN	840
Db	960	FSDVTVTVNVSKEGRSREQFSYVLPVHSLPESMGPKAGGTRITIHGSDNLVNGSMLQVLVN	1019
Qy	841	DTDPCTDLRTATSIITCTVPGGTLSPVPVVCVRFESRGCVHGNLTFWYMQNPVITAIISP	900

Db 1020 DTDPCDLTRTATSITCTVPGGTLPSVPVCVRFESESGCVHGNLTFWYMQNPVITAIISP 1079

Qy 901 RSPVSGGRITITVAGERFHMVQNVSMVAVHHIGREPTFCVKVNLSTLITCPSPGALSNASAPV 960  
|||||

Db 1080 RSPVSGGRITITVAGERFHMVQNVSMVAVHHIGREPTFCVKVNLSTLITCPSPGALSNASAPV 1139  
|||||

Qy 961 DFFINGRAYADEAAEELDPAAEQGRSRFRDLPLNPQFSTAKREKWKHHHGEPLTLVI 1020  
|||||

Db 1140 DFFINGRAYADEAAEELDPAAEQGRSRFRDLPLNPQFSTAKREKWKHHHGEPLTLVI 1199  
|||||

Qy 1021 HKEQDQLGLESHYHIKIGQVSCDIQIISDRVIHCSVNESLGTAEQQLPITIQVGNFNQT 1080  
|||||

Db 1200 HKEQDQLGLESHYHIKIGQVSCDIQIISDRVIHCSVNESLGTAEQQLPITIQVGNFNQT 1259  
|||||

Qy 1081 IATLQLGGSETAIVVSIVICSVLLLLSVVALFVFCTKSRRARIYWKTLQMEEMESQIR 1140  
|||||

Db 1260 IATLQLGGSETAIVVSIVICSVLLLLSVVALFVFCTKSRRARIYWKTLQMEEMESQIR 1319  
|||||

Qy 1141 EEIRKGFELQTDMDTLTKELNRSQGIPFLEYKHVTRTFPPKCSSLYEERYVLPSTLN 1200  
|||||

Db 1320 EEIRKGFELQTDMDTLTKELNRSQGIPFLEYKHVTRTFPPKCSSLYEERYVLPSTLN 1379  
|||||

Qy 1201 SQGGSPQETHPLLGWNIPEHCRPSMEEGISLFSLLNNKHFLIVFVHALEQQKDFAVR 1260  
|||||

Db 1380 SQGGSPQETHPLLGWNIPEHCRPSMEEGISLFSLLNNKHFLIVFVHALEQQKDFAVR 1439  
|||||

Qy 1261 DRCSLASLLTIALHGKLEYYSIMKELLVDLIDASAANKPKMLRRTESVVEKMLTNWMS 1320  
|||||

Db 1440 DRCSLASLLTIALHGKLEYYSIMKELLVDLIDASAANKPKMLRRTESVVEKMLTNWMS 1499  
|||||

Qy 1321 ICMYGLCRETVGEPFFLLCAIKQQINKGSIDAITGKARYTLNEEWLLRENIEAKPRNLN 1380  
|||||

Db 1500 ICMYGLCRETVGEPFFLLCAIKQQINKGSIDAITGKARYTLNEEWLLRENIEAKPRNLN 1559  
|||||

Qy 1381 VSFQCGGMDLSVRAMDTDLTQVKEKILEAFCKNPVYSQWPRAEVDVLEWFASTQSYV 1440  
|||||

Db 1560 VSFQCGGMDLSVRAMDTDLTQVKEKILEAFCKNPVYSQWPRAEVDVLEWFASTQSYV 1619  
|||||

Qy 1441 LRDLDDTSVVEDGRKKLNTLAHYKIPEGASLMSLTDKKDSTLGRVKDLDTEKYFHLVLP 1500  
|||||

Db 1620 LRDLDDTSVVEDGRKKLNTLAHYKIPEGASLMSLTDKKDSTLGRVKDLDTEKYFHLVLP 1679  
|||||

Qy 1501 TDELVEPKKSHRQSHRKKVLPETIYLRLLSTKGTQLKFLDDLKFAILSIREDKPPLAVKY 1560  
|||||

Db 1680 TDELVEPKKSHRQSHRKKVLPETIYLRLLSTKGTQLKFLDDLKFAILSIREDKPPLAVKY 1739  
|||||

Qy 1561 FFDPLEEQAERKGISDPDLHIWKTNSLPLRFVWNILKNPQFVFDIEKTDHIDACLSVIA 1620  
|||||

Db 1740 FFDPLEEQAERKGISDPDLHIWKTNSLPLRFVWNILKNPQFVFDIEKTDHIDACLSVIA 1799  
|||||

Qy 1621 QAFIDACSIDQLGKDSPTNKLLYAKEIPEYRKTVQRYYKQIQDMPPLSEQEMNAHLAE 1680  
|||||

Db 1800 QAFIDACSIDQLGKDSPTNKLLYAKEIPEYRKTVQRYYKQIQDMPPLSEQEMNAHLAE 1859  
|||||

Qy 1681 ESRKYQNEFNTNVAMAEIYKAKRYRPQIMAALEANPTARRTQLQHKFEQVVALMENNIIY 1740  
|||||

Db 1860 ESRKYQNEFNTNVAMAEIYKAKRYRPQIMAALEANPTARRTQLQHKFEQVVALMENNIIY 1919  
|||||

Qy	1741	ECYSEA	1746
Db	1920	ECYSEA	1925

RESULT 4  
 PLXDL1\_HUMAN  
 ID PLXDL1\_HUMAN STANDARD; PRT; 1925 AA.  
 AC Q9Y4D7; Q6PJS9; Q8IZJ2; Q9BTQ2;  
 DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.  
 DT 31-AUG-2004, sequence version 2.  
 DT 25-JUL-2006, entry version 36.  
 DE Plexin-D1 precursor.  
 GN Name=PLXND1; Synonyms=KIAA0620;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
 OC Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=98403880; PubMed=9734811; DOI=10.1093/dnarep/5.3.169;  
 RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,  
 RA Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. X.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro."  
 RL DNA Res. 5:169-176(1998).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND TISSUE SPECIFICITY.  
 RX MEDLINE=22299888; PubMed=12412018; DOI=10.1002/dvdy.10159;  
 RA van der Zwaag B., Hellemans A.J.C.G.M., Leenders W.P.J.,  
 RA Burbach J.P.H., Brunner H.G., Padberg G.W., Van Bokhoven H.;  
 RT "PLEXIN-D1, a novel plexin family member, is expressed in vascular  
 RT endothelium and the central nervous system during mouse  
 RT embryogenesis."  
 RL Dev. Dyn. 225:336-343(2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 1386-1925 (ISOFORMS 1 AND  
 RP 2).  
 RC TISSUE=Muscle, and Uterus;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ussdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP GLYCOSYLATION [LARGE SCALE ANALYSIS] AT ASN-500, AND MASS
RP SPECTROMETRY.
RX PubMed=16335952; DOI=10.1021/pr0502065;
RA Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E.,
RA Moore R.J., Smith R.D.;
RT "Human plasma N-glycoproteome analysis by immunoaffinity subtraction,
RT hydrazide chemistry, and mass spectrometry.";
RL J. Proteome Res. 4:2070-2080(2005).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9Y4D7-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9Y4D7-2; Sequence=VSP_011516;
CC -!- TISSUE SPECIFICITY: Detected at low levels in heart, placenta,
CC lung, skeletal muscle, kidney, thymus and liver. Detected at very
CC low levels in brain, colon, spleen, small intestine and peripheral
CC blood leukocytes.
CC -!- SIMILARITY: Belongs to the plexin family.
CC -!- SIMILARITY: Contains 3 IPT/TIG domains.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AB014520; BAA31595.1; ALT_INIT; mRNA.
DR EMBL; AY116661; AAM49063.1; -; mRNA.
DR EMBL; BC003526; AAH03526.1; -; mRNA.
DR EMBL; BC011848; AAH11848.1; -; mRNA.
DR UniGene; Hs.301685; -.
DR Ensembl; ENSG00000004399; Homo sapiens.
DR HGNC; HGNC:9107; PLXND1.
DR MIM; 604282; gene.
DR InterPro; IPR002909; IPT_TIG_rcpt.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR013548; Plexin_cytopl.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR008936; Rho_GAP.
DR InterPro; IPR001627; Sema.
DR Pfam; PF08337; Plexin_cytopl; 1.
DR Pfam; PF01437; PSI; 2.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF01833; TIG; 3.
DR SMART; SM00429; IPT; 3.
DR SMART; SM00423; PSI; 3.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS51004; SEMA; 1.
KW Alternative splicing; Glycoprotein; Membrane; Polymorphism; Receptor;
KW Repeat; Signal; Transmembrane.
FT SIGNAL 1 46 Potential.
FT CHAIN 47 1925 Plexin-D1.
FT /FTId=PRO_0000024676.
FT TOPO_DOM 47 1271 Extracellular (Potential).
FT TRANSMEM 1272 1292 Potential.
FT TOPO_DOM 1293 1925 Cytoplasmic (Potential).

```

FT	DOMAIN	47	546	Sema.
FT	DOMAIN	891	979	IPT/TIG 1.
FT	DOMAIN	981	1066	IPT/TIG 2.
FT	DOMAIN	1069	1160	IPT/TIG 3.
FT	CARBOHYD	86	86	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	155	155	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	188	188	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	224	224	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	481	481	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	500	500	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	583	583	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	696	696	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	736	736	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	802	802	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	965	965	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1017	1017	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1060	1060	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1099	1099	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1118	1118	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1132	1132	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1237	1237	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1257	1257	N-linked (GlcNAc. . .) (Potential).
FT	VAR_SEQ	1766	1925	SLPLRFWVNILKNPQFVFDIDKTDHIDACLSVIAQAFIDAC SISDLQLGKDSPTNKLLYAKEIPEYRKIVQRYKYQIQDMTP LSEQEMNAHLAEESRKYQNEFNTNVMAEIIYKYAKRYPQI MAALEANPTARRTQLQHKFEQVVALMEDNIIYECYSEA -> RWRPSSPVLGEHPPEEPVCL (in isoform 2). /FTId=VSP_011516.
FT	VARIANT	870	870	M -> V (in dbSNP:2255703). /FTId=VAR_022144.
SQ	SEQUENCE	1925 AA;	212095 MW;	26001F5D0B2A80E5 CRC64;

Query Match 92.2%; Score 8468; DB 1; Length 1925;  
 Best Local Similarity 91.9%; Pred. No. 0;  
 Matches 1606; Conservative 52; Mismatches 88; Indels 2; Gaps 1;

Qy	1	SMLNVAANHNPASTVGLVLPPTS	SGTGGSRLLVGATYTGFGSAFFPRNRSLEDHRFENTPE	60
Db	178	SMLNVAANHNPASTVGLVLP	PAAAGGSRLLVGATYTGYGSSFFPRNRSLEDHRFENTPE	237
Qy	61	IAIRSLDARGDLAKLFTFDLNP	SDDNILKIKQGAKEQHKLGFVRAFLHPAVPPHSAQPYA	120
Db	238	IAIRSLDTRGDLAKLFTFDLNP	SDDNILKIKQGAKEQHKLGFVSAFLHPSDPPGAQSYA	297
Qy	121	YLALNSEARAGDKDSQARSLLARICL	PRGAGGDAKKLTESYIQLGLQCAGGAGRGDLYSR	180
Db	298	YLALNSEARAGDKESQARSLLARICL	PHGAGGDAKKLTESYIQLGLQCAGGAGRGDLYSR	357
Qy	181	LVSVPFAPREQFFAVFERPQ	GAPGARNAPALCAFRFDDVQAAIRAARTACFVEPAPDVVA	240
Db	358	LVSVPFAPRERLFAVFERPQ	GSPAARAAPALCAFRFADVRAIRAARTACFVEPAPDVVA	417
Qy	241	VLDSVVQGTGPACESKRNIQLQPE	QLDCGAHLQHPLTILQPLRASPVFRAPGLTAVAVA	300
Db	418	VLDSVVQGTGPACERKLNILQPE	QLDCGAHLQHPLSILQPLKATPVFRAPGLTSVAVA	477

Qy	361	SANNYTAFLGTATGRLLKISLNESMQVSVRRVLTVAYGEVPHHVMQFDPMDPGYLYLMT	360
Db	478	SVNNYTAFLGTGTNGRLKLNINLESQVSVRRVTVAYGEVPHHVMQFDPADSVYLYLMT	537
Qy	361	SHQMARVKVAACEVHSTCGDCVGAADAYCGWCTLETRCTLQQDCTNSSQPHFWTSASEGP	420
Db	538	SHQMARVKVAACNVHSTCGDCVGAADAYCGWCALETRCTLQQDCTNSSQQHFWTSASEGP	597
Qy	421	SRCPAMTVLPSEIDVHRDYGTMILQISGSLPSLSGSMEMACDYGNGVVRTVARVPGPAYDHQ	480
Db	598	SRCPAMTVLPSEIDVQRQYEPGMILQISGSLPSLSGSMEMACDYGNNIRTVARVPGPAFQHQ	657
Qy	481	IAYCNLLPRAQFPSPFPAQDQHVTVEMSVRVKGHNIVSANFTIYDCSRIGQVYPHTACTSC	540
Db	658	IAYCNLLPRDQFPFPPPNQDQHVTVEMSVRVGRNIVKANFTIYDCSRTAQVYPHTACTSC	717
Qy	541	LSTQWPCSWCIQLHSCSVNSQSQCDSPNPTSPQDCPQILPSPLAPVPTGGSQDILVPLTK	600
Db	718	LSAQWPCFWCSQQHSCSVNSQSRCEASPNPTSPQDCPRTLSPPLAPVPTGGSQNILVPLAN	777
Qy	601	ATFFHGSSLECSFGLEESFEAVWANNLSVRNCQVVLHTTQKSQVFLSLKLKGPDPDRFLD	660
Db	778	TAFFGGAALSCSFGLEEFVAVWNNESVVRCDQVVLHTTRKQSVFLSLQLKGRPARFLD	837
Qy	661	SPNPMTVVVYNCAMGSPDCSQCLGREDLGHLCVWNDGCRRLRGLPLPGTCPAPEIRAIE	720
Db	838	SPEPMTVMVYNCAMGSPDCSQCLGREDLGHLCMSDGCRLRGLPLQPMAGTCPAPEIRAIE	897
Qy	721	PLSGPLDGGTLLTIIRGNLGRRLSDVAHGVMIGSVACEPLADRYTVSEIIVCATGPAAGA	780
Db	898	PLSGPLDGGTLLTIIRGNLGRRLSDVAHGVMIGGVACEPLPDRTVSEIIVCVTGPAAGP	957
Qy	781	FSDVVTVMVNSKEGRSREQFSYVLPTVHSLEPSMGPKAGGTRITIHGSDLNVGSMQLQVLVN	840
Db	958	LSGVVTVMNASKEGSRDRFSYVLPLVHSLEPTMGPKAGGTRITIHGNDLHVGSSELQVLVN	1017
Qy	841	DTDPCDTLRTATSIITCTVPGGTLPSPVVPCVRFESRGCVHGNLTFWYMQNPVITAISPG	900
Db	1018	DTDPCTELMRTDTSIACMTPEGALPAFVPCVRFERRGCVHGNLTFWYMQNPVITAISPR	1077
Qy	901	RSPVSGGRTITVAGERFHMVQNVSMVAVHHIGREPTFCVKVNSTLITCSPGALSNASAPV	960
Db	1078	RSPVSGGRTITVAGERFHMVQNVSMVAVHHIGREPTLCKVNSTLITCSPGALSNASAPV	1137
Qy	961	DFFINGRAYADE--AAEELDPAEAQRGSRFRDYLPNPQFSTAKREKWKIHHHPGEPLTL	1018
Db	1138	DFFINGRAYADEVAAEELDPAEAQRGSRFRDYLPNPQFSTAKREKWKIHHHPGEPLTL	1197
Qy	1019	VIHKEQDSLGLSESHEYHIKIGQVSCDIQISDRVIHCSVNESLGTAEQGLPITIQVGNFN	1078
Db	1198	VIHKEQDSLGLQSHEYRVKIGQVSCDIQIVSDRIHCSVNESLGAAGVGLPITIQVGNFN	1257
Qy	1079	QTIATLQLGGSETAIVSVIVICSVLLLSVVALVFCTKSRAERYWKTLLOMEMESQ	1138
Db	1258	QTIATLQLGGSETAIVSVIVICSVLLLSVVALVFCTKSRAERYWKTLLOMEMESQ	1317



Qy	1139	IREEIIRKGFELQDTMDTLTKELNRSQGIPLFLEYKHVTRTFFPKCSSLYEERYVLPST	1198
Db	1318	IREEIIRKGFELQDTMDTLTKELNRSQGIPLFLEYKHVTRTFFPKCSSLYEERYVLPST	1377
Qy	1199	LNSQGGSPQETHPLLGEWNIPEHCRPSMEEGISLFSLLNNKHFLIVFVHALEQKQDFA	1258
Db	1378	LNSQGGSSQAQETHPLLGEWKIPESCRPNMEEGISLFSLLNNKHFLIVFVHALEQKQDFA	1437
Qy	1259	VRDRCSLASLLTIALHGKLEYYSIMKELLVDLIDASAAKNPKMLRRTESVVEKMLTNW	1318
Db	1438	VRDRCSLASLLTIALHGKLEYYSIMKELLVDLIDASAAKNPKMLRRTESVVEKMLTNW	1497
Qy	1319	MSICMYGCLRETVGEPFFLLCAIKQQINKGSIDAITGKARYTLNEEWLLRENIEAKPRN	1378
Db	1498	MSICMYGCLRETVGEPFFLLCAIKQQINKGSIDAITGKARYTLNEEWLLRENIEAKPRN	1557
Qy	1379	LNVSFQCGCMDSLSVRAMDTDTLTQVKEKILEAFCKNPVYSQWPRAEDVDLEWFASTQS	1438
Db	1558	LNVSFQCGCMDSLSVRAMDTDTLTQVKEKILEAFCKNPVYSQWPRAEDVDLEWFASTQS	1617
Qy	1439	YVLRDLDDTSVVEDGRKKLNTLAHYKIPEGASLAMS LTDKKDSTLGRVKDLDTKEYFHLV	1498
Db	1618	YVLRDLDDTSVVEDGRKKLNTLAHYKIPEGASLAMS LTDKKDSTLGRVKDLDTKEYFHLV	1677
Qy	1499	LPTDELVEPKKSHRQSHRKKVLP E I Y L T R L L S T K G T L Q K F L D D L F K A I L S I R E D K P P L A V	1558
Db	1678	LPTDELAEPKSHRQSHRKKVLP E I Y L T R L L S T K G T L Q K F L D D L F K A I L S I R E D K P P L A V	1737
Qy	1559	KYFFDFLEEQA E K R G I S D P D T L H I W K T N S L P L R F W V N I L K N P Q V F D I E K T D H I D A C L S V	1618
Db	1738	KYFFDFLEEQA E K R G I S D P D T L H I W K T N S L P L R F W V N I L K N P Q V F D I E K T D H I D A C L S V	1797
Qy	1619	IAQAFIDACSI S D L Q L G K D S P T N K L L Y A K E I P E Y R K T V Q R Y Y K I Q D M T P L S E Q E M N A H L	1678
Db	1798	IAQAFIDACSI S D L Q L G K D S P T N K L L Y A K E I P E Y R K T V Q R Y Y K I Q D M T P L S E Q E M N A H L	1857
Qy	1679	A E E S R K Y Q N E F N T N V A M A E I Y K A K R Y R P Q I M A A L E A N P T A R R T L Q L H K F E Q V V A L M E N N	1738
Db	1858	A E E S R K Y Q N E F N T N V A M A E I Y K A K R Y R P Q I M A A L E A N P T A R R T L Q L H K F E Q V V A L M E D N	1917
Qy	1739	IYECYSEA	1746
Db	1918	IYECYSEA	1925